

В.В. Голенков, Н.А. Гулякина, А.В.Губаревич (Минск, БГУИР)

СЕМАНТИЧЕСКАЯ СЕТЬ КАК СПОСОБ ВНУТРЕННЕГО СМЫСЛОВОГО ПРЕДСТАВЛЕНИЯ ИНФОРМАЦИИ В ПАМЯТИ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Многие проблемы искусственного интеллекта будут решены, если мы поймем, что главными объектами формализации в искусственном интеллекте являются не базы знаний интеллектуальных систем и не машины обработки знаний этих интеллектуальных систем, а глобальное смысловое пространство и глобальная интегрированная машина обработки знаний, находящихся в указанном смысловом пространстве.

Без уточнения понятия смысла невозможно решить задачу понимания информации, поступающей на вход интеллектуальной системы по разным каналам, а также невозможно разработать способ внутреннего смыслового представления знаний в памяти интеллектуальной системы, лишенный каких бы то ни было особенностей, обусловленных не смыслом хранимых знаний, а предлагаемой формой их представления.

Смысл (смысловое представление информации) является абстрактной знаковой конструкцией, обладающей следующими свойствами:

(1) Среди знаков, входящих в состав указанной знаковой конструкции не должно существовать пар синонимичных знаков, т.е. знаков имеющих один и тот же денотат, знаков, обозначающих одну и ту же сущность.

(2) Не должно быть дублирования информации — не только в виде многократного вхождения знаков, обозначающих одни и те же сущности (т.е. синонимичные знаки), но и в виде многократного вхождения семантически эквивалентных текстов (знаний), что, собственно, является следствием синонимии знаков.

(3) Среди знаков, входящих в состав указанной знаковой конструкции, не должно существовать омонимичных знаков, т.е. знаков, которые в разных контекстах могут обозначать разные сущности.

(4) Все знаки, входящие в состав указанной знаковой конструкции, не должны иметь внутренней структуры, анализ которой необходим для понимания этой знаковой конструкции. Внутренняя структура знаков необходима только в тех знаковых конструкциях, в которых разрешена множественность вхождения знаков, имеющих один и тот же денотат, и анализ совпадения внутренних структур таких знаков является средством установления их синонимии или предполагаемой синонимии, уточняемой на этапе последующего семантического анализа знаковой конструкции. Следовательно, все знаки, входящие в состав смыслового представления информации, можно считать абстрактными, поскольку они абстрагируются от своего внутреннего устройства, обладают только одним свойством — обозначать взаимно однозначно соответствующие им денотаты, являются инвариантами соответствующих максимальных классов синонимичных знаков.

(5) В состав знаковой конструкции, являющейся смысловым представлением информации, не должно входить ничего, кроме рассмотренных выше абстрактных знаков.

(6) В рамках построения смыслового представления информации не должны использоваться не только такие конструкции, как слова, термины (словосочетания), но и такие языковые приемы, как склонение, спряжение.

(7) С синтаксической точки зрения все знаки, входящие в состав смыслового представления информации должны четко делиться на два вида:

- знаки связей между описываемыми сущностями (связь считается одним из видов описываемых сущностей);
- знаки сущностей, которые связями не являются.

Важным достоинством смыслового представления информации является то, что в нем явно и четко задаются связи между описываемыми сущностями (в т.ч. и связи между связями) в виде связей между знаками этих сущностей: четко указывается семантический тип связи. Подчеркнем также, что все связи каждой описываемой сущности в смысловом представлении информации представлены только множеством знаков связей,

инцидентных знаку указанной сущности, т.е. множеством связей, одним из компонентов которых является знак рассматриваемой сущности. Таким образом, любую знаковую конструкцию можно представить (в семантически эквивалентном виде) как множество знаков описываемых сущностей и множество знаков связей, связывающих эти описываемые сущности с другими сущностями. При этом на описываемые сущности и на связи между ними не накладывается никаких ограничений. Описываемые сущности могут быть:

- материальными (физическими) и абстрактными (виртуальными) — числами, множествами, знаками каких-либо сущностей;
- действительно существующими и вымышленными;
- фиксированными (константными) и произвольными (переменными);
- связями между сущностями.

Нетрудно заметить, что знаковая конструкция, являющаяся смысловым представлением информации, в общем случае не может быть линейной, поскольку каждая описываемая сущность, являющаяся денотатом соответствующего знака, может быть связана неограниченным числом связей с другими сущностями, описываемыми в этой же знаковой конструкции. Таким образом, смысловое представление информации есть графовая структура, имеющая различную теоретико-графовую конфигурацию. Такие графовые структуры, обладающие указанными выше семантическими свойствами, называют также семантическими сетями [1]; [2]. Приведенная трактовка семантических сетей дает возможность в полной мере использовать теорию графов для исследований синтаксических свойств семантических сетей и для построения алгоритмов обработки семантических сетей.

Важнейшим достоинством семантических сетей является то, что в них достаточно просто и наглядно выглядит процедура навигационного поиска знаков описываемых сущностей на основе априорного знания семейства отношений, заданных на этих сущностях.

Следует при этом четко отличать саму семантическую сеть, являющуюся внутренним абстрактным представлением информации в памяти интеллектуальной системы, от ее кодирования в рамках выбранного варианта технической реализации указанной памяти, а также от различных вариантов ее визуализации для пользователей. Сама семантическая сеть и ее графическое или текстовое изображение — не одно и то же. Аналогично этому следует четко отличать графовую структуру как абстрактный объект от различных вариантов ее представления.

Фактически, речь идет о создании формальных средств описания семантики различных видов знаний и формальных средств описания обработки знаний на семантическом уровне. Это предполагает разработку соответствующего стандарта, выделяющего из всего многообразия абстрактных языков семантических сетей определенный базовый универсальный язык семантических сетей, который мы назвали SC-кодом (Semantic Computer code) [3]; [4]; [5]; [6]. Предложенный нами язык является языком

унифицированных семантических сетей. Тексты этого языка будем называть sc-текстами. Знаки, входящие в состав sc-текстов, будем называть sc-элементами. Переход от введенного выше общего понятия семантической сети к унифицированным семантическим сетям (sc-текстам) рассматривается как задание целого ряда ограничений на семантические сети общего вида, но таких ограничений, которые не снижают семантической мощности языка семантических сетей, претендующего на универсальность [7].

Представление смыслового пространства в виде унифицированной семантической сети имеет ключевое значение для решения таких задач, как:

- обеспечение семантической совместимости баз знаний различных интеллектуальных систем;
- интеграция различного вида *знаний*;
- обеспечение компонентного (модульного) проектирования баз знаний интеллектуальных систем путем использования ранее разработанных и апробированных знаний, заимствованных из смыслового пространства.

Смысловое пространство — это результат интеграции всевозможных знаний, представленных в смысловой форме (т.е. в виде семантических сетей). Таким образом, смысловое пространство представляет собой бесконечную семантическую сеть, содержащую всевозможные (и, как минимум, все известные, добытые к текущему моменту) знания. Очевидно, что смысловое пространство должно быть структурировано. Это означает, что в нем должны быть явно выделены и обозначены самые различные знания, а также должны содержаться метазнания, описывающие (специфицирующие) эти выделенные знания.

ЛИТЕРАТУРА

1. Болотова, Л.С. Системы искусственного интеллекта: модели и технологии, основанные на знаниях. — М., 2012.
2. Скороходько, Э.Ф. Семантические сети и автоматическая обработка текста. — Киев, 1983.
3. Голенков В.В. Графодинамические методы и средства параллельной асинхронной переработки информации в интеллектуальных системах. — Минск, 1996.
4. Представление и обработка знаний в графодинамических ассоциативных машинах / В.В.Голенков, О.Е.Елисеева, В.П.Ивашенко и др.; Под ред. В.В.Голенкова. — Минск, 2001.
5. Семантическая модель сложноструктурированных баз данных и баз знаний / В.В. Голенков, Н.А. Гулякина, О.Е. Елисеева и др. — Минск, БГУИР, 2004.
6. Гаврилова Т.А., Голенков В.В., Гулякина Н.А. Унифицированный способ абстрактного кодирования семантических сетей / В сб.: X международная научная конференция имени Т.А.Таран «Интеллектуальный анализ информации ИАИ-2010».: Сб.тр. — Киев, 2010. — С. 46–52.
7. Голенков, В.В. Гулякина, В.В. Структуризация смыслового пространства // Материалы международной научно-техн. конференции «Открытые семантические технологии проектирования интеллектуальных систем» — Минск, 2014. — С. 65–78.